Available online at www.sciencedirect.com

SCIENCE DIRECT

Genomics 87 (2006) 598–607

GENOMICS

www.elsevier.com/locate/ygeno

Comparative genomic organization of the human and bovine *PRNP* locus ☆

Sang-Haeng Choi [a,1], Il-Chul Kim [a,1], Dae-Soo Kim [a], Dae-Won Kim [a], Sung-Hwa Chae [a], Han-Ho Choi [a], Inho Choi [b], Jung-Sou Yeo [b], Mi-Na Song [c], Hong-Seog Park [a,*]

[a] *Genome Research Center, Korea Research Institute of Bioscience and Biotechnology, 52, Oun-dong, Yusong-gu, Daejeon 305-333, Korea*
[b] *School of Biotechnology, Yeungnam University, 214-1, Dae-dong, Gyeongsan, Gyeongsangbuk-do 712-749, Korea*
[c] *Department of Pharmacology and Dental Therapeutics, College of Dentistry, Seoul National University, 28-2 Yeongun-Dong, Chongro-Ku, Seoul, Korea*

Received 4 September 2005; accepted 20 December 2005
Available online 7 February 2006

Abstract

We sequenced a 208-kb BAC clone spanning the bovine prion protein (*PRNP*) locus, and compared the genomic structure with that of human. As a result, we determined the precise breakpoint between the two syntenic genomes, located on the 5′ UTR of the *PRNP* gene, and discovered two highly repetitive sequences near the breakpoint. Further analysis demonstrated that the genomic structure of three genes, *PRNP*, *PRND,* and *RASSF2*, within the syntenic region of the bovine genome is highly conserved in order and orientation. The *PRNT* locus was not found in bovine but is conserved in several primates, including human. Moreover, we confirmed that the bovine *RASSF2* is composed of 10 exons, as is the human gene, showing some difference from a previous report. Our findings may provide useful clues for understanding the evolutional process in the *PRNP* locus and also the mechanism that allows TSE from cattle to infect humans.

Keywords: Prion protein; Transmissible spongiform encephalopathy; Bovine genome; Synteny breakpoint; Organization; Comparative analysis; Repetitive sequence

Introduction

Transmissible spongiform encephalopathy (TSE), also referred to as prion disease, induces several degenerative neurological symptoms, as well as morphological changes characterized by the development of a sponge-like appearance in some areas of the brain. Prion disease is termed Creutzfeld-Jacob disease (CJD) in humans, scrapie in sheep and goats, and bovine spongiform encephalopathy (BSE) in cattle. One of the more serious issues presented by these diseases is that the pathogenic agent that causes the disease is transmissible across species [1]. The most recent prion disease variant, CJD (vCJD), which has recently been detected in humans, is believed to result from the consumption of meat or bone tissue obtained from BSE-infected cattle [2–4].

Many previous efforts have demonstrated that the pathogen is an infectious protein particle called a prion (PrP), and does not contain nucleic acids. The most popular current hypothesis regarding the infection mechanism involves the conversion of the normal cellular PrP (PrP^{C}) into an infectious protein (PrP^{Sc}) via posttranslational processing, in which the PrP β-sheet content is increased, thereby rendering the PrP^{Sc} protein resistant to degradation [5,6].

Furthermore, molecular genetic studies conducted with human subjects have revealed that the *PRNP* gene, as well as other genes involved in the pathogenic mechanism, is localized in the HSA20p12/p13 region [7], which has been designated the *PRNP* locus. Currently, three genes have been mapped to the locus, all located within a 55-kb region (*PRNP*-20 kb-*PRND*-3 kb-*PRNT*) [8–11]. *PRNP* harbors two exons which encode the 287-amino acid PrP protein [12–14]. A variety of insertion mutations within the Cu^{2+}-binding region have also been associated with prion pathogenesis [1]. Doppel (*PRND*), which also harbors two exons, encodes a

☆ DNA sequence data from this article have been deposited with NCBI GeneBank under Accession No. AY944236.

* Corresponding author. Fax: +82 42 879 8139.
E-mail address: hspark@kribb.re.kr (H.-S. Park).

[1] These authors contributed equally to this work.

0888-7543/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.ygeno.2005.12.012

179-amino acid GPI-anchored protein, which bears some similarity to *PRNP,*in terms of both the structure and the topology of the proteins [15]. Doppel shares approximately 25% identity in amino acid sequence with the C-terminal globular domain of the PrP [10,15,16] and is expressed in various tissues during fetal development. The third gene in this locus is *PRNT,* which has recently been identified in humans [11]. *PRNT* encodes three alternative splicing transcripts, and is expressed exclusively in the adult testis. These three genes may be evolutionarily and functionally related.

Understanding the mechanism of TSE pathogenesis requires an understanding of the genomic structure of the *PRNP* locus in several organisms. In addition to humans, much information is currently available for bovines [13,17], pigs [18], mice [10], rats [10], sheep [14], deer [19], goats [20], and zebrafish (Accession No. AI558786). BSE, in particular, has attracted a great deal of public attention, because it is thought that the human prion disease vCJD is caused by exposure to BSE-contaminated bovine tissues. Many studies that have examined the molecular mechanism of TSE have suggested the possibility that not only *PRNP*, but also *PRND*, *PRNT*, and the Ras association domain family 2 (*RASSF2*), are functionally associated with prion disease in humans [8–11]. The genomic structure of these genes has been annotated by the human genome sequencing efforts and they are located within an approximately 150-kb region flanking the *PRNP* gene (NCBI build 35).

However, the comparative genomic structure of the bovine and human loci has not been reported, although the BSE prion protein of cattle is considered to be the source of infection of human vCJD [2–4]. The bovine *PRNP* locus cytogenetically corresponds to the human HSA 20q12/p13 [7]. The D18330 BAC clone (AJ298878), which is 78,056 bp in length, is a unique clone that encompasses the cattle *PRNP* gene, and has been deposited in the NCBI database as a draft sequence [17]. Most studies of the BSE prion have been limited to the identification of single nucleotide polymorphisms (SNPs) of orthologous genes [17,19,21].

To examine the structure of the *PRNP* locus in human and cattle on a genomic scale, we used Korean cattle, which are known to originate from a crossbreeding between the European *Bos primigenius* and the Indian *Bos indicus*, and are known, in the Korean language, as the Hanwoo (Han means Korean and Woo means cattle) strain. In this study, we screened a novel BAC clone from Korean HanWoo cattle, and sequenced 208 kb of the DNA, which encompasses not only the three genes of the human *PRNP* locus, but also the *RASSF2* gene. In this study, we detected a breakpoint in the syntenic region between the loci, and have also presented the comparative genomic organization of three TSE-related genes. Furthermore, we have shown that *PRNT* is absent within the bovine genome, but is conserved to a high degree in primate genomes, and have also described two types of novel repetitive sequences, both of which may be applicable as markers in cattle.

Results

Isolation of a bovine BAC clone homologous to the human PRNP locus

We selected a BAC clone (HW-YUBAC_2) harboring sequences from three exon sequences of the bovine *PRNP* (AJ298878), with a length of 207,929 bp. The only other described clone, D18330 (AJ298878), is a unique BAC clone encompassing a portion of the cattle *PRNP* locus, which is 78,056 bp in length [17]. However, the D18330 clone does not encompass the entire *PRNP* locus of the human counterpart, and the coverage of *PRND* is partial. The BAC clone screened in this study completely encompasses the D18330 clone. The DNA sequences of the HW-BAC clone evaluated in this study were deposited in the NCBI databank, under Accession Number AY944236.

Determining the synteny breakpoint between bovine and human PRNP loci

The synteny block of the *PRNP* locus has been cytogenetically identified as BTA13q17 in bovines, and HAS20q12/p13 in humans [7]; however, a detailed genomic structure, which includes the breakpoint of syntenic regions, has yet to be reported. To determine that position, we selected 230 kb of the genomic sequence flanking the human *PRNP* locus (Accession No. NT_011387). Using the PipMaker program, the human sequence was aligned with 208 kb of the cattle genomic sequence determined in this study, and expressed as a dot-plot sequence comparison graph (Fig. 1). If two queries in a position have 50–100% sequence identity, this is designated with a dot at that position. Via this analysis, we determined that the synteny boundary was 8.7 and 7.1 kb upstream of the 5′ UTR of the *PRNP* gene in bovines and humans, respectively, and further sequence similarities were detected in the 3′ of the *PRNP* gene (Fig. 1a). Approximately 155 kb of the 208-kb HW-BAC clone analyzed in this study corresponded to 140 kb of the human counterpart, but an additional 53 kb of the bovine sequence exhibited no significant sequence similarities to human genome sequences (Fig. 1b).

Identification of two novel repetitive sequences: A 52-bp repeat and the SINE/BovA family

Using the PipMaker program to align the bovine and human sequences, no changes in intragenomic structure, including translocation, inversion, or duplication, were detected. However, two highly repetitive sequences (BLA and BLB) were identified in the nonsyntenic region of the cattle locus (Fig. 1c). The 4201-bp BLA block and the 15,407-bp BLB block were located 33,341 and 17,994 bp from the synteny breakpoint, respectively. Further analysis using the Tandem Repeat Finder ver.3.21 (http://c3.biomath.mssm.edu/trf.html) showed that the BLA repeats were composed of two consensus sequences of 52 bp each (Con A and Con B) showing 87% sequence similarity

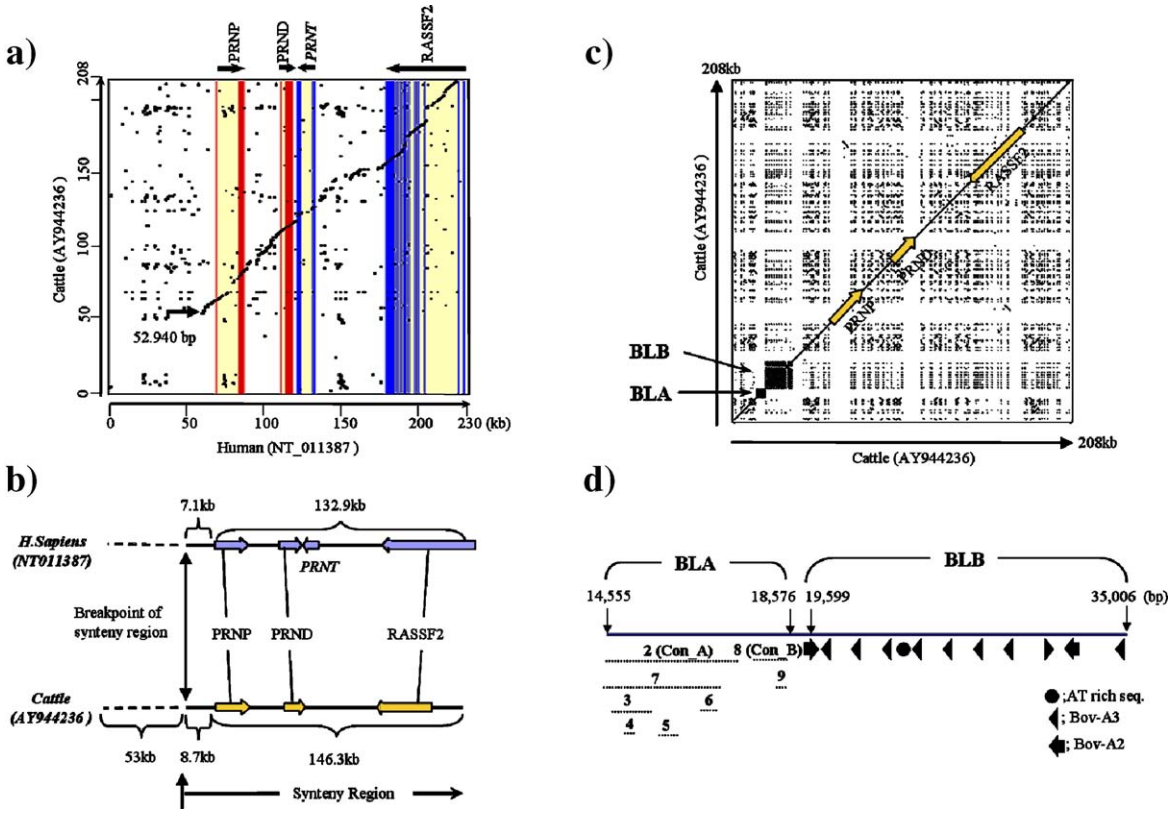

Fig. 1. Dot-plot analysis of genomic organization. (a) Dot-matrix comparison of the human and bovine genomic sequences using the PipMaker program. The percentage similarity is represented by a given dot. Gene names and position within the human locus are represented along the lines. The arrow indicates the boundary of the syntenic region between the bovine and the human *PRNP* loci. Three genes (*PRNP*, *PRND*, and *RASSF2*) exhibited significant sequence similarities, but human *PRNT* did not have a corresponding sequence in cattle. (b) Diagram of comparative genomic organization. A synteny breakpoint position was found 7.1 and 8.7 kb upstream from the 5′ UTR of the human and cattle *PRNP*, respectively. Gene ordering and orientation between the loci were identical, except that the *PRNT* gene is not detected in cattle. (c) Dot-matrix analysis of the intragenomic structure of cattle. A 52-bp tandem repeat sequence (BLA) and the ruminant-specific repeat, SINE/BovA (BLB), were localized in a nonsyntenic region (arrows). (d) Diagram of BLA and BLB repetitive blocks. The BLA block harbors two 52-bp consensus sequences, ConA and ConB, and they were repeated 55.3 and 15.7 times, respectively. The numbers under the dotted lines indicate the index numbers in Table 1. The BLB block contains the SINE/BovA family sequence repeated 11 times. The sequence information of BLA is shown in Table 1.

(Fig. 2). Furthermore, the Con A sequences (14,555-17,462 bp from the synteny breakpoint) were continuously repeated 55.7 times (Table 1 and Fig. 1d), and the Con B sequences (17,780–18,576 bp from the synteny breakpoint) were continuously repeated 15.3 times (Table 1 and Fig. 1d). Additionally, the BLB block was found to be composed of the SINE/BovA family of typical ruminant repeats [22,23], which has been divided into two subfamilies, Bov_A3 and Bov_A2, repeated here 9 and 2 times, respectively (Fig. 1d). Furthermore, our analysis of the bovine *PRNP* locus revealed 26 tandem repeat sequences, excluding simple repeats (Table 1), 2–200 bp in length, repeated 2–56 times, and evidenced an 80–100% sequence similarity among repeat copies.

Gene organization in the bovine PRNP locus

We examined the genomic structure of 140 kb of the human *PRNP* locus (NT_011387) as the counterpart of 155 kb of the bovine sequence, and detected four genes, organized as follows: *PRNP*-20.3 kb-*PRND*-2.8 kb-*PRNT*-39.4 kb-*RASSF2*. However, by way of contrast to what was observed in humans, only three bovine genes were annotated, and were organized as follows: *PRNP*-25.9 kb-*PRND*-43.6 kb-*RASSF2* (Fig. 1, Supplemental Fig. 1). These results indicated that three genes (*PRNP*, *PRND*, and *RASSF2*) are conserved between the two species, and also implied that the homologue of human *PRNT* was absent in bovines, at least at the *PRNP* locus (Supplemental Table 2, Fig. 3). The comparative gene structures for each of the genes are described below.

PRNP

Bovine *PRNP* was characterized initially via long-PCR amplification [24], and the genomic structure was analyzed using a BAC clone (AJ298878) [25].

Here, we confirmed that the bovine *PRNP* gene contains three exons (53, 98, and 4,091 bp in length, respectively) separated by two introns (2436 and 13,551 bp in length), as compared to two exons (90 and 2,384 bp) and one intron (12,698 bp) in humans. The ORF length of the bovine gene was 795 bp, in contrast to 762 bp in humans. The sequence identity was 83.0% at the nucleotide level, and 86.0% at the amino acid level. We detected minor differences in the

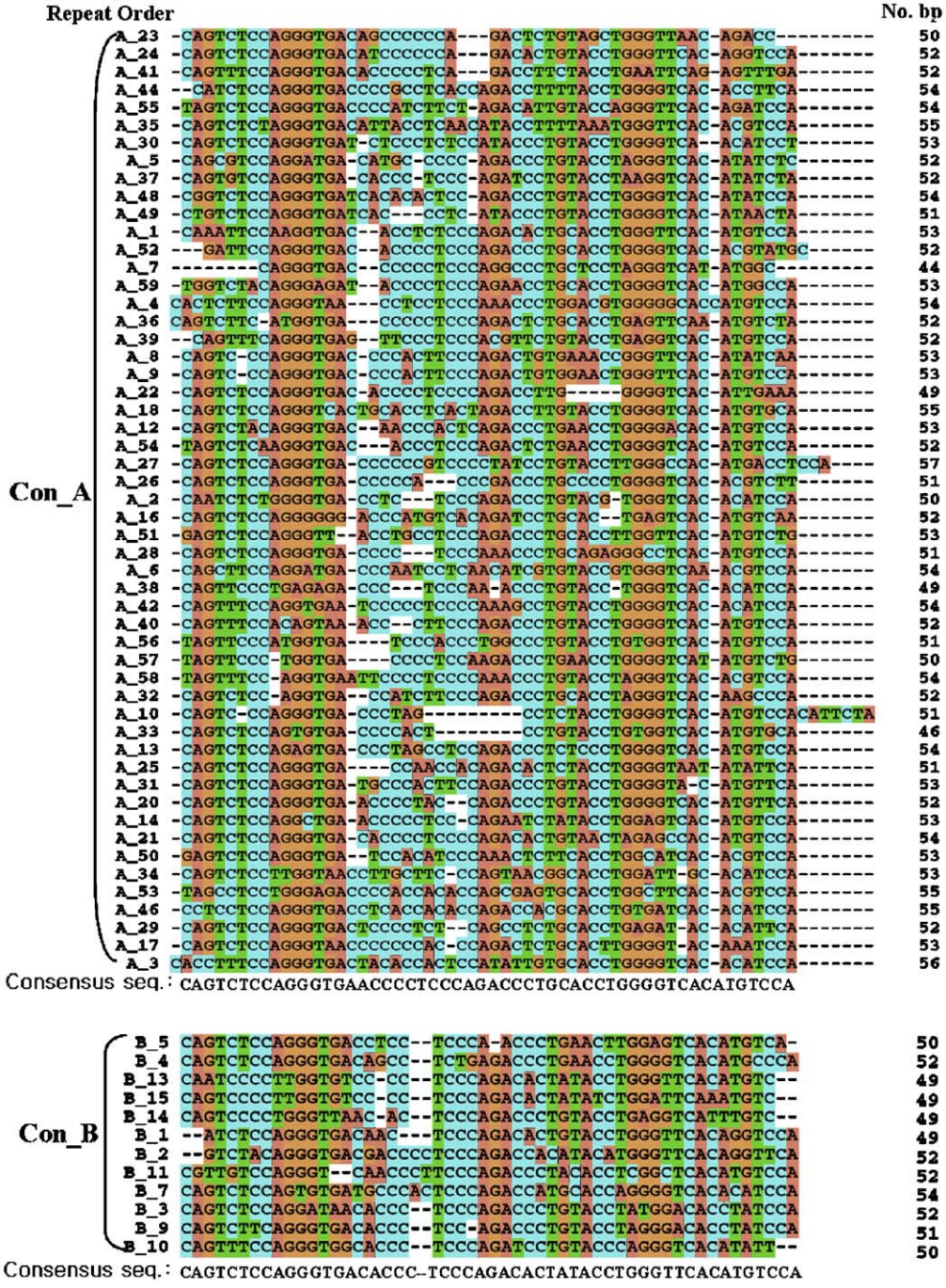

Fig. 2. Sequence alignment of the BLA block. The sequences of the 52-bp repeat units are clearly assembled into two subgroups, Con_A and Con_B, via sequence identity. The consensus sequence similarity between Con_A and Con_B was approximately 87%. The numbers of repeated sequences were 55.7 for Con_A and 15.3 times for Con_B.

intragenic sequences (data not shown), but the overall genomic structure was consistent with the findings of previous reports [14,24,25].

PRND

The comparative structure of the human and bovine *PRND* gene has been reported by Comincini et al. [17], using long-PCR for the bovine sequence (AY017310), and was recently updated by the same group (AY017310.4). We found that the *PRND* gene contained two exons in bovines, measuring 89 and 3219 bp, which were different from Comincini's result [17], but were consistent with what was observed in the sequences of humans (60 and 3922 bp) (NM_012409 and NC_000020) and mice (54 and 3122 bp) (NM_023043, NC_000068). Furthermore, coding sequences (CDS) begin in the second exon in both human and cattle, and the bovine sequence (537 bp) is 6 bp longer than that of humans (531

Table 1
Tandem repeat ristribution in 208-kb genome sequence flanking PRNP gene in the Korean cattle Hanwoo

Indices	Consensus size	Copy number	Percent matches	Consensus sequence	Index
11648–11673	5	5.2	100	ACTGA	(1)
14555–17462	52	55.7	65	CAGTCTCCAGGGTGAACCCCTCCCAGACCCTGCACCTGGGGTCACATGTCCA	(2)
14696–15171	104	4.5	74	CCCCCTCCCAGAACCTGCACCTGGGGTCACATGGCCACAGTCTCCAGGGTGAACCCCTACCAGACCCTGTACCTGGGGTCACATGTTCAG AGTCTCCAGGGTTA	(3)
14995–15141	53	2.7	78	TCCAGGGTGACCCCCCTCCCAGACCCTGCACCTGGGGTCACACCTATGCCATC	(4)
15281–15529	51	4.8	74	CAGACCCTGTACCTGGGGTCACATATCCACAGTCTCCAGGGTGACCAACCA	(5)
16691–16898	52	4	79	CCCAGACCCTGAACCTGGGGTCACACGTCCACAGTCTCCAGGGTGACCCCCA	(6)
14418–16937	103	24.1	67	CAGTCTCCAGGGTGACCCCCACCCAGACCCTGCACCTGGGGTCACACGTCTCAGTCTCCAGGGTGACACCCTCCCAGACACTGAACCTGG GGTCACATGTCCA	(7)
17780–18576	52	15.3	68	CAGTCTCCAGGGTGACACCCTCCCAGACACTATACCTGGGTTCACATGTCCA	(8)
18320–18576	103	2.5	89	CCCCTCCCAGACACTATACCTGGATTCAAATGTCCACAGTCTCCAGGATGACACCCTCCCAGACCCTGTACCTAGGGACACCTATCT ACAATCCCCTTGGTGT	(9)
30502–30547	23	2	95	CAGAACCCTCCCCACTCAACTCT	(10)
36171–36293	41	3	97	TACCAAGAGGAATTTCACTTGAAATGAAAGGGCCTTTACTT	(11)
56772–6796	2	12.5	100	AC	(12)
67164–67197	17	2	100	TAAATATTTATTTAATT	(13)
68118–68149	16	2	100	TTCTCCTTCCTGAGTT	(14)
78033–78157	24	5.2	93	GGTGGCTGGGGCCAGCCCCATGGA	(15)
80302–80330	14	2.1	100	TTTTTGTAAGGTAC	(16)
84757–84781	5	5	100	ATCAG	(17)
87831–87871	17	2.4	95	GGGGATCTTCTGGACCC	(18)
91761–91785	5	5	100	CTGAA	(19)
99791–99823	16	2.1	94	TGTGGCCCCATGGACA	(20)
103860–104096	119	2	85	AGAATACTGGAATGGGTAGCCGTTCCCTTCTCCAGGAGATCATCCCAACCCAGGGAACAAACCCAGGTCTCCCGCACTGCACGTGGAT TCTTTACCAGCTGAGATACCAGGGAAGCCCA	(21)
115906–116149	119	2	80	AATTCTTGGGCTTCCCTGATAGCTCAGCTGGTAAAGAAACCACCTGCAATGCTGGAGACCCTGGTTCAATTCCTGGGCCAGGAAGA TCTCCTGGAGAAGGGAAAAGCTACCCACTCCAG	(22)
133883–134255	186	2	94	ATTTCCTTCTCCAATGCATGAAAGGGAAAAGTCAAAGTGAAGTCACTCAGTCGTGTCCAACCCTTAGTGACCCCATGGACTACAGCCTACCAG GCTCCTCCATCCATGGGATTTTCCAGGCAAGAGTACTGGAGTAGAGTGCCATTGCCTCCTCCATTGCTTCTACTAATCCATGACAATTTGAAT	(23)
147925–147983	29	2	100	AATGAATGGTTCCCTCCTCCCCCAAGAGC	(24)
194227–194263	17	2.1	90	AGATCCCTGAGTAGGAA	(25)
201288–201313	5	5.2	100	TTCAG	(26)

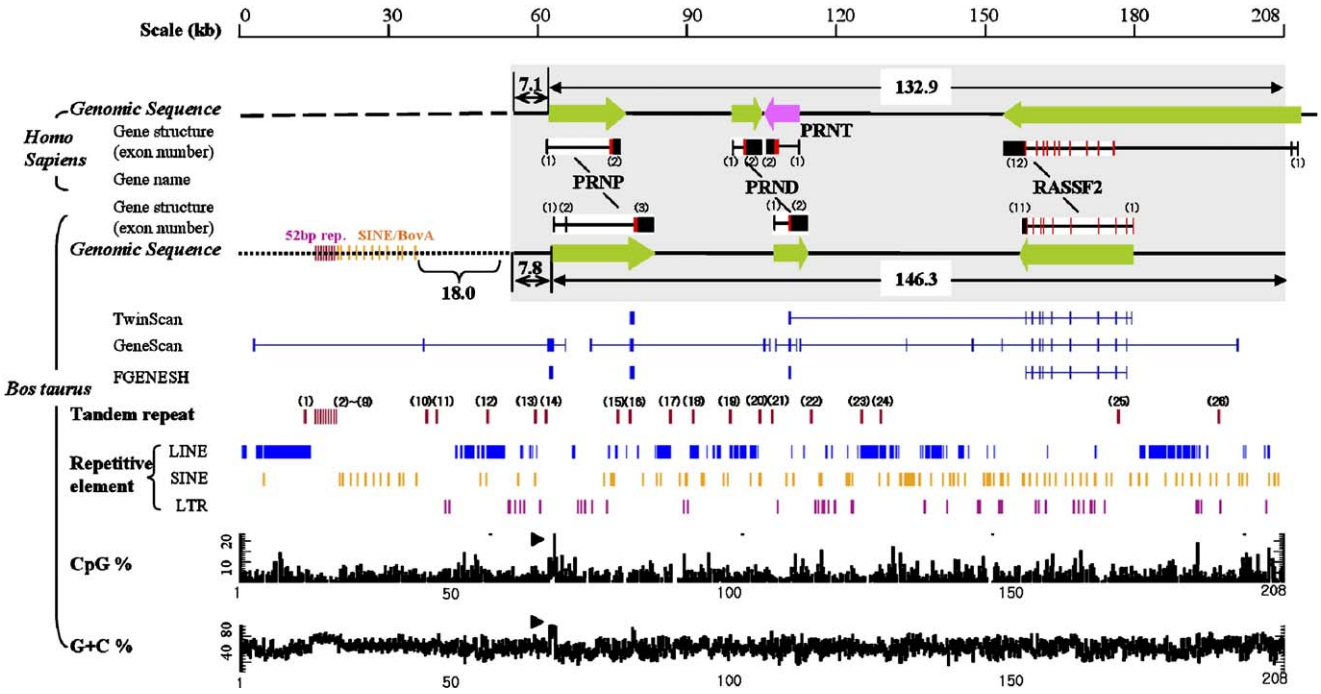

Fig. 3. Summary of comparative genomic organization of the *PRNP* locus between cattle (208 kb) and human (230 kb, NT_011387). The syntenic and nonsyntenic regions are represented by a gray box and a dotted line, respectively. The synteny breakpoints in humans and cattle were localized to 7.1 and 7.8 kb upstream of the 5′ UTR of *PRNP*. Two repetitive sequences in the bovine genome, 52 bp and SINE/BovA family, are located 18.0 kb from the synteny breakpoint. The ordering and orientation of the three orthologous genes are represented by the green box, and the gene structures are represented above (cattle) or below (human) the gene symbol. The pink box represents the human *PRNT* gene, which is deleted in bovines. A vertical blue bar within the gene structure shows the CDS region of each gene. Gene structures predicted using the TwinScan, GeneScan, and FGENESH programs are represented by blue lines. The CDS for the three orthologous genes were predicted precisely by all three programs. The positions of 26 tandem repeats and repetitive elements, including SINE, LINE, and LTR, are represented by vertical bars. Orphan CpG islands indicating the presence of cryptic genes were found in the *PRNP* region (arrowheads). Unfortunately, the genomic sequences of the bovine BAC clone determined in this study did not encompass the first and second exons of human *RASSF2* (NT_011387), contributing to the lesser size of bovine *RASSF2* in comparison to humans.

bp). Finally, the sequence identity between the genes was 81.0% at the nucleotide level and 75.0% at the amino acid level.

RASSF2

RASSF2 in cattle was first characterized by Comincini's group [26], using exon trapping and RACE. They determined that the gene harbored 9 exons, measuring 1807 bp in total (AF277231), in comparison to the 10 exons found in humans (KIAA0168). However, our study detected 10 exons (1832 bp), identical to what was detected in human *RASSF2*. This difference was revealed by sequence alignment, conducted using the SIM4 program (data not shown), and we also detected a 7-bp insertion (TTTAAAA) in the second exon of Comincini's sequence. The sequence identity existing between our sequence and Comincini's was 99.8% at the nucleotide level and 100% at the amino acid level. Despite the high level of sequence identity, the 7 bp discrepancy between the two sequences critically affected exon selection in the transcript. The CDS derived from our sequences contained 10 exons, comprising 981 bp in length, whereas Cominicini's harbored 9 exons and 987 bp (Fig. 4, Supplemental Fig. 1).

To validate this 7 bp discrepancy, we amplified the genomic sequences from 10 individual HanWoo, and from another four strains of cattle (Simmental, Limousine, Holstein, and Brown Swiss), using primers encompassing the 7-bp region. The resulting sequence alignment revealed that this 7 bp was inserted only in Cominicini's sequence, AF277231 (data not shown). In accordance with these results, we concluded that this discrepancy is caused not by strain specificity between HanWoo and *Bos taurus* (BBI-B759KO0741) but as the result of a sequence error in the previous study [26]. However, we cannot completely rule out the possibility of strain specificity as the cause of the 7 bp discrepancy, due to our small sample size.

To compare the genomic structures of bovine and human *RASSF2*, we extracted two distinct ORFs (NP_739580, NP_055552, and NP_739579) and three alternative cDNA sequences (NM_170774, NM_014737, and NM_170773) from the NCBI databank. When the bovine *RASSF2* was compared with these human sequences, we determined that the human NM_170774 sequence had the highest structural similarity, although the two were considerably different in length (5262 bp in the human vs 1832 bp in the bovine (Supplemental Table 2). Further analysis showed that this was the result of the 3′ UTR length difference—819 bp in bovines, and 4381 bp in humans. However, both ORFs were identical in length (981 bp), and 8 of 10 exons corresponded completely in length, with only the first and second exons

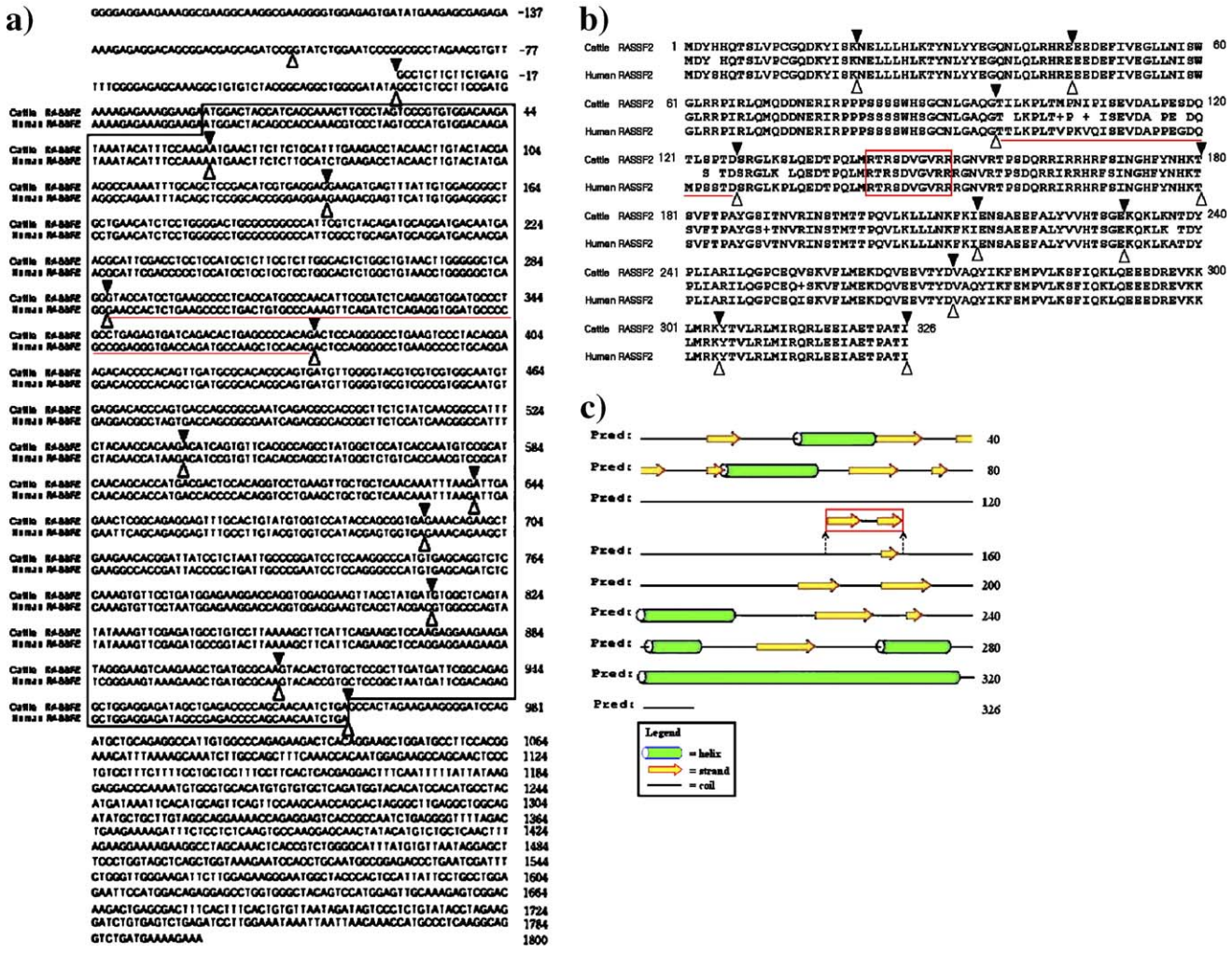

Fig. 4. Structure comparison of human and bovine *RASSF2*. (a) Comparative DNA sequence analysis. The exon positions are represented by white triangles for humans, and black triangles for the cattle sequences. The CDS region is shown with a black box. The sequence identities between exons were approximately 90%, with the exception of the fourth exon (85%), which was represented by a red underline. (b) Comparative analysis of amino acid sequence. Sequence identities in 9 of 10 exons ranged from 90 to 100%, but were only 68% in the fourth exon, underlined in red. (c) Comparison of the secondary structure of the *RASSF2* protein using the PSIPRED program. Prominent changes were detected between amino acids 142 and 150, and these alterations predicted one β-sheet in the bovine, but two β-sheets in the human protein.

differing (Fig. 4a). The degree of sequence identity between human and cattle ORFs was 90.5% at the nucleotide level and 94.2% at the amino acid level (Fig. 4b). In addition, we attempted to determine the predicted secondary structure of the *RASSF2* protein, using the PSIPRED program (http://bioinf.cs.ucl.ac.uk/psipred/) with a CDS of 326 amino acids. Both of the ORFs generated almost identical structures, with the exception of a 9-amino acid difference (aa 142–150), which resulted in the generation of one β-sheet in bovines and two β-sheets in humans (Fig. 4c).

PRNT is absent in bovine but present in primate

The first characterization of the *PRNT*gene (NM_177549.2) from the human genome revealed that this gene harbors two exons, which can give rise to three transcripts [11]. We conducted a comparison of the genomic structure of the HW-BAC sequence (AY944236) determined in this study with the human genome sequence (NT_011387), using the PipMaker and Sim4 programs. However, we were unable to detect any bovine sequences corresponding to human *PRNT,* and we confirmed this using three gene prediction programs, FGENE, GENESCAN, and TWINSCAN (Supplemental Fig. 1). According to the results of an analysis using BLASTN and BLASTX with the NCBI and SWISS-PROT data bases, we detected a genomic sequence on chimpanzee chromosome 21 that matched human *PRNT*, showing a 98% identity in nucleotide sequence (data not shown). These data suggest a possibility that the *PRNT* gene is highly conserved in several primates.

To address this possibility, we amplified genomic DNAs obtained from chimpanzees, gorillas, orangutans, and crab-eating macaques (*Macaca fascicularis*), using primers designed from the CDS of the human *PRNT* gene. Interestingly, all of the

primate genomic DNAs amplified a 285-bp fragment, identically to humans, but no PCR products were detected in the cattle samples (data not shown).

Discussion

We screened a novel BAC clone that was localized on the *PRNP* locus in Korean HanWoo cattle, and identified 208 kb of a DNA sequence. Employing several publicly available bioinformatics tools, we conducted a comparative analysis of genomic structures between human and bovine loci.

The breakpoint of the homologous genomic region between the loci was approximately 8.7 kb, and was 7.1 kb upstream from the 5′ UTR of the *PRNP* gene in bovines and humans, respectively. We also determined that the syntenic regions between bovines and humans continued downstream of the *PRNP* gene via the *RASSF2* gene, with a length of approximately 155 kb in bovines and 140 kb in humans. Since the BAC clone employed in this study was shorter than that of the human counterpart, we were unable to determine the synteny break position of the 3′ end of the locus. In addition, the genome size of the syntenic regions differed by 15 kb, due to differences in the LINE frequency between the two loci (Supplemental Table 1).

The bovine *PRNP* locus has been studied previously [17,25–27], but the genomic structure has not yet been defined. In this study, we found that the gene orientation within the bovine *PRNP* locus was *PRNP*-25.85 kb-*PRND*-43.55 kb-*RASSF2*. *PRNP*, *PRND,* and *RASSF2* are highly conserved in the bovine genome but *PRNT* was not found in the bovine genome. A further analysis using gene prediction software such as GRAILEXP, GENESCAN, and TWINSCAN confirmed these findings. Although a previous report [11] identified nucleotide sequence similarity between the human *PRNT*and genomic sequence of mouse (68%) and the ESTs of rat (55–68%) and zebrafish (59%), we could not find significantly conserved genomic sequence in either the *PRNP* locus of the mouse (NT_039207) or the rat (NW_047658) or elsewhere in the databanks (NCBI/EMBL/DDBJ) that corresponds to the human *PRNP*. This result means that the *PRNT* gene is present in humans but not in rodents, which agrees with the conclusion reported by Marko et al. [28]. Moreover, we found that a unique genomic sequence which showed 98.6% nucleotide identity to the human *PRNT* ORF was found on the chimpanzee chromosome 21 (UCSC browser), and we reconfirmed by PCR amplification methods that the CDS sequence is conserved in several primates, including human as well as chimpanzee. These results strongly indicate that the *PRNT* gene appeared in the primate lineage after the evolutionary split from rodents.

Furthermore, recent studies have suggested that the *PRNT* gene may play a role in the development of prion disease in humans [11,29]. However, our results indicated that *PRNT* does not exist in bovines, the animal in which TSE most frequently occurs. Therefore, it could be a possible interpretation that *PRNT* may not play an important role in the development of TSE, or the mechanism underlying TSE in cattle is different from that in humans. Further research will be required to elucidate the functional role of *PRNT* and its role in prion diseases.

Previous studies have indicated that bovine *RASSF2* possesses a completely different gene organization from humans, although the two species evidence a high degree of sequence identity in the CDS [26], whereas the gene structures of *PRNP* and *PRND* are conserved to a high degree [10,14,25,27]. We confirmed the structures of *PRNP* and *PRND* and defined their genomic organization. In addition, we determined that bovine *RASSF2*, as seen with *PRNP* and *PRND*, is highly conserved in terms of genomic organization, as compared to humans. Not only do both genes in bovines and humans contain 10 exons, but also, all but the first and second exons correspond completely in terms of length. Furthermore, the predicted secondary structures of the bovine and human *RASSF2* proteins harbored a different β-sheet number at amino acids 142–150. Rasfadin (*RASSF2*) is believed to be one of the prion-mediated downstream signaling components [26,30,31].

The variable repeat sequences detected in the human genome are believed to function in interchromosomal rearrangements [32], and are shared only by closely related species [33]. In this study, we detected two novel sequence blocks that harbored repetitive sequences, including the ruminant-specific repeat of the SINE/BovA family and a 52-bp repetitive sequence unit, which are tandemly repeated 11 times and 71 times, respectively. Moreover, they are located proximally to the boundary of synteny between human and cattle. These repetitive sequences imply that the nonsyntenic region performed an important function in gene shuffling by chromosome recombination, and the novel repetitive sequences identified in this study should provide useful clues in our understanding of the evolution of bovine and closely related species.

Materials and methods

BAC library screening

The BAC library of Korean HanWoo cattle was kindly provided by Dr. I. Choi at the Yeungnam University of Korea (unpublished). We screened the BAC clone carrying *PRNP,* using PCR amplifications of the three exons of cattle *PRNP*. The three pairs of PCR primers were designed as follows: *PRNP*-E1L (5′-AAACATTGGGCTCAATCCAG-3′) and *PRNP*-E1R (5′-CAACGAG-CACGCTCAACAG-3′) for exon 1, *PRNP*-E2L (5′-AAAGCTGCATTCACC-CACAG-3′) and *PRNP*-E2R (5′-GTGATCCCAGCCTTTCAGTC-3′) for exon 2, and *PRNP*-E3L (5′-GCAACCGTTATCCACCTCAG-3′) and *PRNP*-E3R (5′-TGGCTTACTGGGTTTGTTCC-3′) for exon 3. PCR cycling conditions were as follows: denaturation at 96°C for 20 s, annealing at 56°C for 20 s, and extension at 72°C for 3 min, with 35 amplification cycles.

Construction of a random plasmid library

The selected BAC clones were grown overnight in 100 ml LB medium with 25 mg/ml of chloramphenicol at 37°C with shaking. The BAC clone DNA was extracted and purified using a Plasmid Midi Kit (Qiagen, Hilden, Germany). To construct the random plasmid library for shotgun sequencing, the extracted BAC clone DNA (10–15 μg) was randomly sheared by DNA hydroshear (Genomic Solutions Corp. Ann Arbor, MI), and subsequently fractionated (0.7–1.5 kb)

with Chroma Spin+TE–1000 Columns (BD Biosciences Clontech, Mountain View, CA). The fractionated DNAs were dephosphorylated and ligated into pUC118 vector using Takara BKL (Blunting Kination Ligation) Kit (Takara Bio Inc., Otsu, Japan). The plasmid DNAs to be transformed in DH-5α-competent cells were extracted and purified via alkaline lysis, using 96-well filter plates (AcroPrep 96; Pall Corp., NY).

DNA sequencing

The prepared DNAs were sequenced directly with an internal sequencing primer, using the BigDye Terminator (Ver. 3.1) cycle sequencing kit (Applied Biosystems, Foster City, CA). Cycle sequencing was conducted using a GeneAmp PCR System 9700 (Applied Biosystems) using 3 μl of PCR cocktail (250 ng plasmid DNA, 0.5 μl 3 pmol primer, 0.87 μl 5× buffer, 1.38 μl distilled water, and 0.25 μl BigDye). Amplification was conducted with 35 cycles of denaturation at 96°C for 10 s, annealing at 50°C for 5 s, and extension at 60°C for 4 min. The PCR products were then purified via ethanol precipitation, and resolved on an ABI 3730XL DNA analyzer (Applied Biosystems). The individual sequences were assembled with Phred/Phrap software (Ewing and Green, 1998; Ewing et al., 1998; www.phrap.org), the value of the Phred scores of the sequences was 40 or higher, and the average depth of coverage was at least 8×.

The closing of remaining gaps was achieved via direct sequencing using plasmid DNA. Sequence editing for consensus contig formation was generated by visual confirmation, using the Sequencher 4.1.4 program (Gene Codes Corp., Ann Arbor, MI). The sequences reported in this paper have been deposited in the GenBank database, under Accession Number AY944236.

Amplification of a PRNT fragment

To examine the *PRNT* gene, PCR primers were designed from the CDS sequences of the human *PRNT* gene (NM_177549.2), which was registered in NCBI as follows: *PRNT*-CDSF (5′-ATGCAGCATAGTCTTGTTTTC-3′) and *PRNT*-CDSR (5′-CAAAAAAGATTACTGCTAA-3′). Genomic DNAs as a template were used at a concentration of 10–20 ng/ μl for the amplification of the *PRNT* gene. The PCR cycling conditions were as follows: 35 cycles of denaturation at 96°C for 20 s, annealing at 55°C for 20 s, and extension at 72°C for 3 min. Primate genomic DNA samples (chimpanzee, gorilla, orangutan, and crab-eating macaque; *M. fascicularis*) were kindly provided by Dr. Fujiyama of RIKEN GSC, Japan.

Genomic structure analysis of the PRNP locus

The genomic structure analysis of the *PRNP* locus was conducted with several analysis programs; RepeatMasker for the identification of genome-wide repeats and low complexity regions (Jurka et al., 1996; http://ftp.genome.washington.edu/RM/RepeatMasker.html), the CpG plot program for the plotting of CpG island distribution (Larsen et al., 1992), Tandem Repeat Finder (Ver 3.21) for the analysis of tandem repeats (http://c3.biomath.mssm.edu/trf.html), BLAST programs for searching sequence homology in the NCBI database (Altschul et al., 1990; http://www.ncbi.nlm.nih.gov/blast), SIM4 for the alignment of the mRNA sequence with the genomic sequence (Florea et al., 1998), and PipMaker for the alignment of the HW BAC clone sequence and the human genomic sequence as its counterpart (Schwartz et al., 2000; http://bio.cse.psu.edu/). Further searches for potential genes on the contig sequences used GRAILEXP (Uberbacher and Mural, 1991), GENESCAN (http://genome.dkfz-heidelberg.de/cgi-bin/GENSCAN/genscan.cgi), and TWINSCAN (http://genes.cs.wustl.edu/).

Acknowledgments

This work was supported, in part, by a BioGreen 21 grant from the Ministry of Agriculture and by the Ministry of Science and Technology of Korea. The authors are grateful to Dr. T. Taylor for the review of the paper, Dr. Y. Kuroki (RIKEN GSC, Japan) for the provision of sample supplies, and all of the technical staff (Genome Structure Research Laboratory, KRIBB of Korea) for their excellent technical assistance.

Appendix A. Supplementary data

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.ygeno.2005.12.012.

References

[1] J. Collinge, Prion diseases of humans and animals: their causes and molecular basis, Annu. Rev. Neurosci. 24 (2001) 519–550.
[2] M.E. Bruce, R.G. Will, J.W. Ironside, I. McConnell, D. Drummond, A. Suttie, L. McCardle, A. Chree, J. Hope, C. Birkett, S. Cousens, H. Fraser, C.J. Bostock, Transmissions to mice indicate that 'new variant' CJD is caused by the BSE agent, Nature 389 (1997) 498–501.
[3] M.R. Scott, R. Will, J. Ironside, H.O. Nguyen, P. Tremblay, S.J. DeArmond, S.B. Prusiner, Compelling transgenetic evidence for transmission of bovine spongiform encephalopathy prions to humans, Proc. Natl. Acad. Sci. USA 96 (1999) 15137–15142.
[4] E.A. Asante, J.M. Linehan, M. Desbruslais, S. Joiner, I. Gowland, A.L. Wood, J. Welch, A.F. Hill, S.E. Lloyd, J.D. Wadsworth, J. Collinge, BSE prions propagate as either variant CJD-like or sporadic CJD-like prion strains in transgenic mice expressing human prion protein, EMBO J. 21 (2002) 6358–6366.
[5] S.B. Prusiner, Novel proteinaceous infectious particles cause scrapie, Science 216 (1982) 136–144.
[6] S.B. Prusiner, Prions, Proc. Natl. Acad. Sci. USA 95 (1998) 13363–13383.
[7] B. Castiglioni, S. Comincini, B. Drisaldi, T. Motta, L. Ferretti, Comparative mapping of the prion gene (*PRNP*) locus in cattle, sheep and human with PCR-generated probes, Mamm. Genome 9 (1998) 853–855.
[8] N.K. Robakis, E.A. Devine-Gage, E.C. Jenkins, R.J. Kascsak, W.T. Brown, M.S. Krawczun, W.P. Silverman, Localization of a human gene homologous to the PrP gene on the p arm of chromosome 20 and detection of PrP-related antigens in normal human brain, Biochem. Biophys. Res. Commun. 140 (1986) 758–765.
[9] R.S. Sparkes, M. Simon, V.H. Cohn, R.E. Fournier, J. Lem, I. Klisak, C. Heinzmann, C. Blatt, M. Lucero, T. Mohandas, et al., Assignment of the human and mouse prion protein genes to homologous chromosomes, Proc. Natl. Acad. Sci. USA 83 (1986) 7358–7362.
[10] R.C. Moore, I.Y. Lee, G.L. Silverman, P.M. Harrison, R. Strome, C. Heinrich, A. Karunaratne, S.H. Pasternak, M.A. Chishti, Y. Liang, P. Mastrangelo, K. Wang, A.F. Smit, S. Katamine, G.A. Carlson, F.E. Cohen, S.B. Prusiner, D.W. Melton, P. Tremblay, L.E. Hood, D. Westaway, Ataxia in prion protein (PrP)-deficient mice is associated with upregulation of the novel PrP-like protein doppel, J. Mol. Biol. 292 (1999) 797–817.
[11] E. Makrinou, J. Collinge, M. Antoniou, Genomic characterization of the human prion protein (PrP) gene locus, Mamm. Genome 13 (2002) 696–703.
[12] H.A. Kretzschmar, L.E. Stowring, D. Westaway, W.H. Stubblebine, S.B. Prusiner, S.J. Dearmond, Molecular cloning of a human prion protein cDNA, Dna 5 (1986) 315–324.
[13] C. Puckett, P. Concannon, C. Casey, L. Hood, Genomic structure of the human prion protein gene, Am. J. Hum. Genet 49 (1991) 320–329.
[14] I.Y. Lee, D. Westaway, A.F. Smit, K. Wang, J. Seto, L. Chen, C. Acharya, M. Ankener, D. Baskin, C. Cooper, H. Yao, S.B. Prusiner, L.E. Hood, Complete genomic sequence and analysis of the prion protein gene region from three mammalian species, Genome Res. 8 (1998) 1022–1037.
[15] H.P. Mo, R.C. Moore, F.E. Cohen, D. Westaway, S.B. Prusiner, P.E. Wright, H.J. Dyson, Two different neurodegenerative diseases caused by

proteins with similar structures, Proc. Natl. Acad. Sci. USA 98 (2001) 2352–2357.
[16] G.L. Silverman, K. Qin, R.C. Moore, Y. Yang, P. Mastrangelo, P. Tremblay, S.B. Prusiner, F.E. Cohen, D. Westaway, Doppel is an N-glycosylated, glycosylphosphatidylinositol-anchored protein. Expression in testis and ectopic production in the brains of Prnp(0/0) mice predisposed to Purkinje cell loss, J. Biol. Chem. 275 (2000) 26834–26841.
[17] S. Comincini, M.G. Foti, M.A. Tranulis, D. Hills, G. DiGuardo, G. Vaccari, J.L. Williams, I. Harbitz, L. Ferretti, Genomic organization, comparative analysis, and genetic polymorphisms of the bovine and ovine prion Doppel genes (*PRND*), Mamm. Genome 12 (2001) 729–733.
[18] T. Martin, S. Hughes, K. Hughes, M. Dawson, Direct sequencing of PCR amplified pig PrP genes, Biochim. Biophys. Acta 1270 (1995) 211–214.
[19] M.P. Heaton, K.A. Leymaster, B.A. Freking, D.A. Hawk, T.P. Smith, J.W. Keele, W.M. Snelling, J.M. Fox, C.G. Chitko-McKown, W.W. Laegreid, Prion gene sequence variation within diverse groups of U.S. sheep, beef cattle, and deer, Mamm. Genome 14 (2003) 765–777.
[20] C. Billinis, C.H. Panagiotidis, V. Psychas, S. Argyroudis, A. Nicolaou, S. Leontides, O. Papadopoulos, T. Sklaviadis, Prion protein gene polymorphisms in natural goat scrapie, J. Gen. Virol. 83 (2002) 713–721.
[21] D. Hills, J. Schlaepfer, S. Comincini, I. MacLean, G. Dolf, L. Ferretti, I. Olsaker, J.L. Williams, Sequence variation in the bovine and ovine *PRNP* genes, Anim. Genet. 34 (2003) 183–190.
[22] F.G. Sheikh, S.S. Mukhopadhyay, P. Gupta, Pstl repeat: a family of short interspersed nucleotide element (SINE)-like sequences in the genomes of cattle, goat, and buffalo, Genome 45 (1) (2002) 44–50.
[23] J.A. Walker, D.A. Hughes, B.A. Anders, J. Shewale, S.K. Sinha, M.A. Batzer, Quantitative intra-short interspersed element PCR for species-specific DNA identification, Anal. Biochem. 316 (2003) 259–269.
[24] M. Horiuchi, N. Ishiguro, H. Nagasawa, Y. Toyoda, M. Shinagawa, Genomic structure of the bovine PrP gene and complete nucleotide sequence of bovine PrP cDNA, Anim. Genet. 29 (1998) 37–40.
[25] D. Hills, S. Comincini, J. Schlaepfer, G. Dolf, L. Ferretti, J.L. Williams, Complete genomic sequence of the bovine prion gene (*PRNP*) and polymorphism in its promoter region, Anim. Genet. 32 (2001) 231–232.
[26] S. Comincini, B.M. Castiglioni, G.M. Foti, I. Del Vecchio, L. Ferretti, Isolation and molecular characterization of rasfadin, a novel gene in the vicinity of the bovine prion gene, Mamm. Genome 12 (2001) 150–156.
[27] M.A. Tranulis, A. Espenes, S. Comincini, G. Skretting, I. Harbitz, The PrP-like protein Doppel gene in sheep and cattle: cDNA sequence and expression, Mamm. Genome 12 (2001) 376–379.
[28] M. Premzl, J.E. Gready, L.S. Jermiin, T. Simonic, J.M. Graves, Evolution of vertebrate genes related to prion and shadoo proteins-Clues from comparative genomic analysis, Mol. Biol. Evol. 21 (12) (2004) 2210–2231.
[29] S. Preuss, T. Peischl, E. Melchinger, H. Geldermann, Numerous polymorphic microsatellites in the human prion gene complex (including *PRNP*, *PRND* and *PRNT*), Gene 329 (2004) 197–203.
[30] W.C. Shyu, M.C. Kao, W.Y. Chou, Y.D. Hsu, B.W. Soong, Heat shock modulates prion protein expression in human NT-2 cells, Neuroreport 11 (2000) 771–774.
[31] J. Satoh, Y. Kuroda, S. Katamine, Gene expression profile in prion protein-deficient fibroblasts in culture, Am. J. Pathol. 157 (2000) 59–68.
[32] R.F. Wintle, T.G. Nygaard, J.A. Herbrick, K. Kvaloy, D.W. Cox, Genetic polymorphism and recombination in the subtelomeric region of chromosome 14q, Genomics 40 (1997) 409–414.
[33] K. Krejci, J. Koch, An in situ study of variant telomeric repeats in human chromosomes, Genomics 58 (1999) 202–206.